AMENDMENTS TO THE SPECIFICATION

Kindly amend paragraph [0019] as follows:

[0019] FIG. 1 provides the basic modules that are used in a spoken dialog system 100. A user 102 that is interacting with the system will speak a question or statement. An automatic speech recognition (ASR) module 104 will receive and process the sound from the speech. The speech is recognized and converted into text. AT&T's Watson ASR component is an example of such an ASR module. The text is transmitted to a spoken language understanding (SLU) module 106 (or natural language understanding (NLU) module) that determines the meaning of the speech, or determines the user's intent in the speech. This involves interpretation as well as decision: interpreting what task the caller wants performed and determining whether there is clearly a single, unambiguous task the caller is requesting - or, if not, determining actions that can be taken to resolve the ambiguity. The NLU 106 uses its language models to interpret what the caller said. The NLU processes the spoken language input wherein the concepts and other extracted data are transmitted (preferably in XML code) from the NLU 106 to the [[DM]] dialog manager (DM) application 108 along with a confidence score. The (DM) module 108 processes the received candidate intents or purposes of the user's speech and generates an appropriate response. In this regard, the DM 108 manages interaction with the caller, deciding how the system will respond to the caller. This is preferably a joint process of the DM engine 108 running on a Natural Language Services (NLS) platform (such as AT&T's infrastructure for NL services, for example) and the specific DM application 108 that it has loaded and launched. The DM engine 108 manages dialog with the caller by applying the compiled concepts returned from the NLU 106 to the logic models provided by the DM application 108. This determines how the system interacts with a caller, within the context of an ongoing dialog. The substance of the response is transmitted to a spoken language generation component (SLG) 110 which generates

words to be spoken to the caller 102. The words are transmitted to a text-to-speech module 112 that synthesizes audible speech that the user 102 receives and hears. The SLG 110 either plays back pre-recorded prompts or real-time synthesized text-to-speech (TTS). AT&T's Natural Voices ® TTS engine provides an example of a TTS engine that is preferably used. Various types of data and rules 114 are employed in the training and run-time operation of each of these components.